Orthographic effects in Mandarin spoken language production

Qingqing Qu[1 2], Markus F. Damian[3]

[1]Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China

[2]Department of Psychology, University of Chinese Academy of Sciences, Beijing, China

[3]School of Psychological Science, University of Bristol, United Kingdom

Word count: 6,575

Address for correspondence:

Qingqing Qu

Key Laboratory of Behavioral Science

Institute of Psychology

Chinese Academy of Sciences, Beijing, China

16 Lincui Road, Chaoyang District, Beijing, China

100101

Tel: +86-10-64888629

Fax: +86-10-64872010

Email: quqq@psych.ac.cn

Abstract

For literate individuals, does the spoken production of language involve access to orthographic codes? Previous research has rendered mixed results, with a few positive findings contrasting with a range of null findings. In the current experiments, we chose spoken Mandarin as the target language in order to better dissociate sound from spelling. Mandarin speakers named coloured line drawings of common objects with adjective-noun phrases (e.g., /lan2/ /hua1ping2/, "blue vase"). Adjectives and nouns were semantically and phonologically unrelated on all trials, but on critical trials they shared an orthographic radical. In two experiments, this resulted in a significant facilitatory effect on naming latencies. We interpret these results as strong evidence for the claim that retrieval of phonological codes in spoken production involves the co-activation of orthographic representations.

Keywords: language production; speaking; orthography; interactivity; Mandarin Chinese

1. Introduction

In a seminal article, Seidenberg and Tanenhaus (1979) first documented an orthographic influence on spoken word identification: when participants performed timed rhyme judgement on spoken word pairs, rhyming pairs which were similarly spelled (pie-tie) were judged faster than pairs for which spelling differed (rye-tie). Since then, effects of spelling in spoken word identification have been documented in a wide variety of tasks (e.g., Castles, Holmes, Neath, & Kinoshita, 2003; Chéreau, Gaskell & Dumay, 2007; Dijkstra, Roelofs, & Fieuws, 1995; Goswami, Ziegler, & Richardson, 2005; Morais, Cary, Alegria, & Bertelson, 1979; Pattamadilok, Perre, Dufau, & Ziegler, 2009; Perre & Ziegler, 2008; Perre, Pattamadilok, Montant, & Ziegler, 2009; Zecker, Tanenhaus, Alderman, & Siqueland, 1986, and others; but see Cutler & Davis, 2012; Mitterer & Reinisch, 2015, for contrary evidence).This may imply that listeners cross-activate orthographic codes *on-line* whenever a spoken word is processed, via bidirectional functional links between orthography and phonology (e.g., Chéreau et al., 2007; Pattamadilok et al. 2009). Alternatively, effect of spelling could emerge *off-line,* resulting from the restructuring of phonological representations during literacy acquisition (e.g., Montant, Schön, Anton & Ziegler, 2011; Perre et al., 2009). This restructuring view assumes that the nature of phonological representations is altered during the process of learning to read and write, leading to “phonographic” representations that integrate orthographic knowledge. Note that on- and offline accounts are not mutually exclusive.

From the abundance of findings on spelling effects in receptive tasks, one might predict parallel effects in spoken production. However, the available evidence (see Table 1 for an overview) is more limited and less consistent.

As can be appreciated from this overview, these results paint an inconsistent picture. Approximately half of the studies utilised the “implicit priming” task, a popular tool in research on spoken word production. Speakers repeatedly produce a small set of spoken responses within experimental blocks, and form overlap within a block is manipulated. Word-initial phonological overlap between responses within a block results in a facilitatory effect on naming latencies (e.g., Meyer, 1990, 1991), interpreted as the online use of partially available information for speech planning (Roelofs, 1997). With English speakers, Damian and Bowers (2003) reported orthographic effects: priming arising from shared word-initial segments (“coffee”, “camel”, “climate”) was disrupted if one item was substituted with one which had a conflicting word-initial grapheme (“coffee”, “camel”, “kennel”). However, this positive result contrasts with several null findings across various languages (Dutch: Roelofs, 2006; French: Alario, Perre, Castel & Ziegler, 2007; Mandarin: Bi, Wei, Janssen & Han, 2009; Zhang & Damian, 2012), suggesting that if orthographic effects are genuine, the implicit priming task does not reliably detect them.[1]

By contrast, a number of recent contributions have demonstrated orthographic effects in spoken word production in tasks which required the learning of novel words. Rastle, McCormick, Bayliss and Davis (2011) asked participants to associate novel spoken words with novel objects, and only subsequently introduced the spelling of the novel words. Target objects were associated with words such that their initial phonemes could be either spelled regularly or irregularly based on English spelling–sound relationships (e.g., /kIsp/ spelled as *kisp* or *chisp*). Orthographic regularity effects were obtained on naming latencies of the novel objects, with faster responses to regular items than irregular ones (as well as in perception tasks; however, not in auditory shadowing), and the authors argued that object naming involves simultaneous activation of phonological and orthographic codes. Bürki, Spinelli and Gaskell (2012) studied the acquisition of novel French words containing consonant clusters which can be pronounced either with or without a schwa, although the reduced variants are more frequent in speech. For example, the initial consonant cluster of the novel French word /pluR/ typically results in a schwa reduction. Bürki et al. asked speakers to associate, over several days, the spoken reduced variants of these novel words with novel objects, and then introduced spelling of the novel words which either did, or did not, orthographically represent the schwa (i.e., “pelour” vs. “plour”). In a subsequent naming test, speakers produced more schwa variants for words which had been spelled accordingly than for words which had not. Moreover, novel words with an orthographic representation of the schwa were produced slower compared to those which did not orthographically represent the schwa. The slower responses were taken to be the consequence of competition between the phonological representation of the reduced variants established by repeated auditory exposure, and the phonological representation for the schwa variant generated via the novel word’s spelling. The authors favoured an offline account, according to which orthographic exposure changes the way in which phonological variants are stored and processed. Similarly, Han and Choi (2015) used a similar word learning technique to explore allophonic variants of /h/ in Korean, and again obtained effects of orthography, interpreted as offline restructuring of phonological codes via their spelling.

Results from a range of additional tasks (see Table 1) have rendered inconsistent results.[2 3] In evaluating the available evidence, it is of course wise to treat null findings with caution. Nonetheless, it should be clear that the results exhibit considerable inconsistency. Perhaps this reflects the difference between online and offline effects, as recently argued by Bürki et al. (2012), with production tasks exclusively detecting offline processes. Some of the used tasks might be of questionable ecological validity, such as the implicit priming task which requires speakers to produce the same few responses over and over again. The inconsistency may also be explained by the fact that researchers of spoken production are generally quite restricted in their choice of materials. Finally, in alphabetic scripts, spelling and sound are strongly confounded, hence it is difficult to design experiments in which the two dimensions are properly dissociated.

To make progress, one would hence ideally use a task which a) is plausibly sensitive to potential online interactions of sound and spelling, b) has at least some degree of ecological validity, and c) involves a target language in which spelling and sound can be largely dissociated (i.e., a non-alphabetic orthographic system). In the experiments reported below, we provided this sort of evidence. In Experiment 1, Mandarin speakers named coloured objects with adjective-noun phrases, and on critical trials, adjective and noun were orthographically related, which resulted in a significant facilitation effect. In Experiment 2, we replicated this pattern with slightly modified materials, and modified various aspects of the design in an attempt to reduce the likelihood of strategic variables impacting the effect. The orthographic effect from the first experiment was again found. These results convincingly demonstrate the online activation of orthographic codes in spoken production.

2. Experiment 1

2.1 Method

2.1.1. Participants. Twenty-seven native speakers of Mandarin Chinese (15 females, mean age 22 years, range 20-28 years) participated in the experiment and were paid RMB35 (approximately US$5). None were colour blind, and all had normal or corrected-to-normal vision and no history of neurological or language problems. The project was approved by the Ethics Committee of the Institute of Psychology, Chinese Academy of Sciences.

2.1.2 Materials and Design. Four colours (blue, brown, green, and orange) and 12 line drawings of objects with no canonical colour from the Snodgrass and Vanderwart (1980) picture set were used. All colour names in Chinese were monosyllabic and all picture names were disyllabic. The average lexical frequency of object names was 3.6 per million in the Chinese Lexicon (Chinese Linguistic Data Consortium, 2003) database, and they could be written with an average of 16 strokes. Each colour was combined with 3 objects to form 12 *orthographically related* colour-object pairings (e.g., 蓝花瓶, blue vase, /lan2hua1ping2/; 橙梳子, orange comb, /cheng2shu1zi /; the two words shared an orthographic radical[4]). Colours and objects within the orthographically related condition were then recombined to form the unrelated condition with 12 orthographically unrelated pairings (e.g., 橙花瓶, orange vase, /cheng2hua1ping2/; 蓝梳子, blue comb, /lan2shu1zi/). In this way, identical stimuli were used across the two conditions (see Appendix A). Care was taken to minimise semantic or phonological overlap (in terms of onset, rhyme, and tone). We collected semantic ratings on a seven-point Likert scale for all colour-object combinations from a group of 16 native Chinese participants (1 = “not related at all”, and 7 = “closely related”). Average rating scores were 2.8 and 2.9 for orthographically related and unrelated combinations, respectively ($t < 1$). Hence, stimuli were semantically well matched across orthographically related and unrelated combinations. As in English, adjectives precede nouns in Chinese.

Each participant was presented with 6 blocks of 24 trials with each of the related and unrelated combinations appearing once in each block, for a total of 144 trials. A new pseudo-random order of trials was generated for each participant and block, such that neither pictures nor colours were repeated on consecutive trials.

2.1.3. Procedure. Stimuli were presented using E-Prime 1.1 software (Psychology Software Tools, Pittsburgh, PA). Naming latencies were measured using a voice-key, connected with the computer via a PST Serial Response Box. Participants were first asked to familiarise themselves with the experimental stimuli by viewing them on the screen, with the expected name printed underneath each object. Subsequently, participants were told that they would see the objects in different colours presented on a computer screen, and their task was to name them with an adjective-noun combination as quickly and accurately as possible, e.g., 蓝椅子, /lan2yi3zi/, ‘blue chair’. Next, participants received a practice block consisting of 8 objects which were not from the set of targets (each of the four colours was repeated twice). Subsequently, the six experimental blocks were presented, separated by a short break. On each trial, participants saw a fixation cross (500ms), a blank screen (500ms), and a picture which disappeared once the participant initiated a verbal response, or after 4,000ms. The intertrial interval was 2,000ms. The experimental session lasted approximately 20 minutes.

2.2 Results and Discussion

Latencies on trials with incorrect responses (0.3%) and faster than 200ms or slower than 1,800ms (3.0%) were excluded. Average response times and errors are presented in Table 2, showing a facilitatory effect (16ms) of orthographic overlap on latencies.

Latencies were analysed using a linear mixed-effects model (Baayen, Davidson, & Bates, 2008; Bates, 2005). Preliminary data analysis showed that there was considerable variability among items in their naming times, with variability arising not only from the object, but also (and in fact more so) from the colour adjective (e.g., items in “blue” are named 100 ms faster than the average, and “orange” 121 ms slower than average). In other words, variability of latencies for objects was confounded with variability of latencies for colours. In order to partial out the variance associated with colours, *colour* was included as a fixed effect in all analyses which by itself is not of great interest. We initially constructed a “maximal model” (Barr, Levy Scheepers & Tily, 2013) which contained the fixed factors *relatedness* and *colour*, as well as adjustments to intercepts and slopes for the random effects *participants* and *object names*. However, the model showed clear evidence of over-parameterisation via r = 1.00 (perfect correlations between intercept and slope adjustments for *object names*) and such a complex random effect structure is therefore not appropriate (Baayen, Davidson & Bates, 2008). When the random effect structure was stepwise reduced, the “most complex" model which did not suffer from overparameterisation included slope adjustments only for *participants*, and intercept adjustments for *participants* and *object names*. The comparison of the "most complex" model with the "maximal model" was not significant, suggesting that removing random slopes adjustments for *participants* did not reduce the fit, $\chi 2(N = 3{,}758) = 1.41$, $p = .49$. Critically, the most complex model showed a significant effect of *relatedness*, $\beta = 16$, $SE = 7.37$, $F = 4.80$, $p = .038$, and colour, $F = 178$, $p < .001$.[5]

Parallel analyses were conducted on the errors but with a binomial family due to the binary nature of the data (Jaeger, 2008). All models which included slope adjustments to *participants*, and *object names* showed evidence of overparameterisation. The most complex model was therefore the one which included intercepts only as random effects. In this model, *relatedness* was not significant, $\beta = -0.96$, $SE = 0.79$, *Wald z* = -1.22, $p = .220$.

In summary, the results showed a significant facilitatory effect on latencies when colour and object name shared an orthographic radical. However, there were some limitations in Experiment 1. First, although in choosing our stimuli we had attempted to avoid semantic or phonological overlap between colour and object names, there were two combinations in the orthographically related condition (绿线轴, green cotton reel, /lü4xian4zhou2/; 棕松鼠, brown squirrel, /zong1song1shu3/), and one combination in the orthographically unrelated condition (棕苍蝇, brown fly, /zong1cang1ying/), in which colour and first syllable of object names had matching tone. Second, it has been suggested that colour prototypicality of objects affects naming times (Naor-Raz, Tarr, & Kersten, 2003), such that objects in typical colours (yellow banana) are named faster than objects in atypical colours (purple banana). While we generally avoided objects with highly prototypical colours in our Materials, one of the related combinations ("brown squirrel") was potentially problematic in this regard. Third, stimuli on half of the trials were orthographically related, and the same unrelated/related combinations were shown six times across the experiment, which might have directed participants' attention to the orthographic manipulation. And finally, because participants were familiarised with the object names prior to the experiment (see "Procedure"; a standard practice in experiments on spoken word production) they were explicitly exposed to orthographic properties of the target words.

In Experiment 2, apart from attempting to replicate the central finding, we aimed to extend Experiment 1 in the following ways. First, we used a revised set of materials in which additional care was taken to avoid residual phonological overlap or colour-object association. Second, to discourage potential strategies, we added a further 12 filler pictures to reduce the percentage of related trials to 25%, and we reduced the number of repetitions of each related/unrelated combination from six to three. Moreover, in the familiarisation phase, we introduced object and colour names verbally to participants. Finally, we conducted postexperimental interviews and asked participants after each testing session whether they had noticed a relationship between colour and object name.

3. Experiment 2

3.1 Method

3.1.1. Participants. Thirty-two native Chinese speakers (20 females, mean age 22 years, age range 21-27 years), none of whom had been in the first experiment, participated in this experiment, and were paid RMB 35 (approximately US$5).

3.1.2 Materials, Design and Procedure. All aspects of Experiment 2 were the same as those of Experiment 1, except that (1) we used a slightly revised set of materials, with an average lexical frequency of 3.41 per million (Chinese Linguistic Data Consortium, 2003) and an average stroke number of 15; see Appendix B; (2) besides 12 critical pictures, a further 12 filler pictures were added in order to reduce the percentage of related trials. As was the case for the critical target pictures, each filler picture was paired with two of the critical colours, thus forming 24 filler trials in which each colour appeared six times. Semantic, phonological or orthographic overlap between adjective and noun was minimized. Each combination was repeated three times, thus generating 144 trials in total, presented in three blocks of 12 related, 12 unrelated, and 24 filler trials; (3) expected names of the pictures and colours were not presented visually in the familiarisation stage, but instead the experimenter named them to participants; (4) after testing, participants were asked to report whether they had noticed any relation between colour and picture names.

The experiment was run using DMDX (Forster & Forster, 2003), and vocal responses were recorded using a microphone connected to the computer. Vocal responses were inspected and analysed by a research assistance who was blind to the hypotheses and design of the study using CheckVocal (Protopapas, 2007).

3.2 Results and Discussion

Data were analysed in the same way as described in Experiment 1. Latencies on trials with incorrect responses (5.3%) and faster than 200ms or slower than 1,800ms (0.7%) were excluded. Filler trials were not analysed. Average response times and errors are shown in Table 2, showing a facilitatory effect (39ms) of orthographic overlap.

Analysis of latencies with a mixed-effects model showed that as for the results from Experiment 1, the maximal model with the full random effect structures was over-parameterised (perfect correlations between intercept and slope adjustments for *participants*). The simplified model included slope adjustments only for *object names*, and intercept adjustments for *participants* and *object names*. The comparison of the simplified model with "maximal model" was not significant, suggesting that removing random slopes adjustments did not reduce the fit, $\chi 2(N = 2{,}166) = 2.81$, $p = .25$. The simplified but most complex model showed a significant effect of *relatedness*, $\beta = 38$, $SE = 14.6$, $F = 6.84$, $p = .03$, and *colours*, $F = 27.97$, $p < .001$.

Parallel analyses were conducted on the errors but with a binomial family due to the binary nature of the data (Jaeger, 2008). All models which included slope adjustments to *participants* and/or *object names* showed evidence of overparameterisation. The most complex model was therefore the one which only included random intercepts. In this model, *relatedness* was not significant, $\beta = -0.12$, $SE = 0.19$, *Wald* $z < 1$, $p = .537$.

Postexperimental interviews revealed that none of the participants had recognised the orthographic relation between colour and object names.

4. General Discussion

Current evidence on whether for literate individuals, the preparation of spoken language is affected by orthographic properties is mixed and inconsistent (e.g., Alario, Perre, Castel & Ziegler, 2007; Damian & Bowers, 2003; see Table 1 for details). In two experiments we presented strong evidence for such an involvement of orthography: when Mandarin speakers named coloured objects via adjective-noun phrases, orthographic overlap between the two words induced a facilitatory effect. The coloured object naming task is a well-established tool in research on spoken production and has been used both with speakers of Western languages (e.g., Damian & Dumay, 2009), and Mandarin (Qu, Damian & Kazanina, 2012). The task has at least superficial ecological validity and the current results are unlikely to have arisen from strategies that participants developed.

We acknowledge that the size of the orthographic effect varied considerably across the two experiments, for reasons not yet determined. Differences in materials are unlikely to be relevant, as stimuli overlapped to a large extent between the two studies. Interestingly, the second experiment in which additional unrelated filler items had been inserted showed a *larger* orthographic effect than the first experiment, further arguing against a strategic origin of the effect. Participants also named objects faster in the second (853 ms) compared to the first (978 ms) experiment, with the larger orthographic effect arising in the faster experiment. A speed-accuracy tradeoff is a possibility, with Experiment 2 showing faster response latencies, but higher error rates (5.4%) than Experiment 1 (0.3%). It remains to be determined whether in the coloured picture naming task, the size of the orthographic effect varies with overall speed, or depends on some other property of materials or participants.

Our results underscore the usefulness of non-alphabetic languages in order to investigate a potential role of orthography in speaking: because sound and spelling are largely independent, experiments can vary orthographic properties (in this case, radical overlap) while avoiding phonological overlap. Equivalent experiments in languages with alphabetic orthography are difficult, although not impossible, to implement (see e.g., Roux & Bonin, 2011, in which orthography and phonology were manipulated independently in a *written* picture naming task so that French target and context pictures shared the initial letter but not the initial sound, as in “cigar-camion”, or they shared the initial phonemes but not the initial letter, as in “singe-ceinture”). Given our evidence supporting a role of orthography in speaking in the current experiments, how can one account for the considerable degree of inconsistency in previous findings (see Introduction, and Table 1)? Of course, some of the failures to obtain effects in spelling might simply be false negatives. Nonetheless, some null findings are obtained fairly consistently, e.g., in the “implicit priming” task for which one positive finding is countered by six negative findings. What could account for such null findings if orthographic effects in other tasks are accepted as genuine? As briefly summarised in the Introduction, a general distinction is between “online” and “offline” sources of potential effects, with the former reflecting direct processing crosstalk between spelling and sound, whereas the latter attributes effects of orthography to a restructuring of phonology during literacy acquisition. Bürki et al. (2012) suggested that there is no “online” crosstalk between spelling and sound in production tasks; rather, to the extent that orthographic effects in production arise, they reflect “offline” restructuring of phonological representations during acquisition of literacy. According to Bürki et al., this could account for the positive findings on word learning tasks, but the null findings on “implicit priming” tasks (see Table 1). However, the results from the current experiments do quite clearly reflect “online” cross-activation between spelling and sound, so the offline/online distinction favored by Bürki et al. appears less relevant.

A different possibility is that phonological and orthographic codes are accessed at different speeds in different tasks, with orthographic effects only emerging in tasks in which orthographic codes are accessed simultaneously with, or perhaps even slightly ahead of, access to phonological codes. This possibility was discussed by Rastle et al. (2012, p. 1592) in order to account for their positive findings of orthography in picture naming, but a null finding in auditory shadowing. In the shadowing task, participants hear a spoken word and are instructed to immediately repeat it, thus phonological activation can guide responses before activation of orthographic codes can exert an influence. By contrast, in picture naming, a task in which responses are much slower than in shadowing, orthographic and phonological representations are activated simultaneously, hence there is opportunity for orthographic feedback to influence spoken responses. This account however appears somewhat ad hoc when jointly considering all existing evidence from Table 1. It is acknowledged that direct evidence on the time course of access to phonological vs orthographic codes in spoken production, perhaps via EEG, would be extremely useful (see Zhang & Damian, 2009, for an initial attempt).

Our findings demonstrate activation of orthographic codes during phonological encoding, arising from orthographic (i.e., radical) overlap within a two-word spoken phrase. How could a mechanism which explains this finding be integrated into computational accounts of word production (Dell, 1986; Dell, Schwartz, Martin et al., 1997; Levelt, Roelofs & Meyer, 1999)? Naming a coloured object requires phonological encoding of both adjective and picture name. Orthographic forms of adjective and noun could be either directly activated from meaning (as they presumably would in written picture naming), or alternatively, phonological encoding of the response might result in co-activation of corresponding orthographic forms via bilateral links. Under either scenario, orthographically related colour-object pairs would prime each other at the orthographic level, and activation would be required to be transmitted, via bilateral links, to the phonological level, resulting in a priming effect in naming latencies. Note that such an account does not necessitate sub-lexical correspondences between sound and spelling, which in non-alphabetic languages such as Mandarin Chinese (our target language) are obviously much reduced compared to alphabetic languages. In the domain of language perception (rather than production), frameworks exist which incorporate such bilateral links between orthography and phonology. For instance, the Bimodal Interactive Activation Model (BIAM; Grainger & Ferrand, 1994) implements visual and spoken word recognition via orthographic and phonological pathways which are bidirectionally connected both at the sublexical and the lexical level. A recent extension (Diependaele, Ziegler & Grainger, 2010) additionally implements an output phonological layer. To adapt this architecture to semantically driven language production, one would need to add higher-level representations (conceptual; lexical-semantic; lexical-syntactic) which activate output phonology. Via cross-activation of the implemented orthographic and phonological pathways, such a model could plausibly account for orthographic effects such as those shown here. In summary, our experiments provide evidence for a genuine orthographic effect in spoken phrase production by Mandarin speakers. The results eventually will need to be accounted for in a computational framework of spoken production which implements online cross-talk between phonological and orthographic representations.

References

Alario, F.-X., Perre, L., Castel, C., & Ziegler, J. C. (2007). The role of orthography in speech production revisited. *Cognition, 102,* B464–B475.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390–412.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255-278.

Bates, D. (2005). Fitting linear mixed models in R. *R News*, 5(1), Article 5. Retrieved from http://ftp.cs.pu.edu.tw/network/CRAN/ doc/Rnews/Rnews_2005-1.pdf

Bi, Y. C., Xu, Y. D., & Caramazza, A. (2009). Orthographic and phonological effects in the picture-word interference paradigm: Evidence from a logographic language. *Applied Psycholinguistics, 30*, 637–658.

Bi, Y., Wei, T., Janssen, N., & Han, Z. (2009). The contribution of orthography to spoken word production: Evidence from Mandarin Chinese. *Psychonomic Bulletin and Review, 16*, 555–560.

Biedermann, B. & Nickels, L. (2008). Homographic and heterographic homophones in speech production: Does orthography matter? *Cortex, 44*(6), 683-697.

Bürki, A., Spinelli, E., & Gaskell, G. (2012). A written word is worth a thousand spoken words: The influence of spelling on spoken-word production. *Journal of Memory and Language, 67,* 449–467.

Castles, A., Holmes, V. M., Neath, J., & Kinoshita, S. (2003). How does orthographic knowledge influence performance on phonological awareness tasks? *Quarterly Journal of Experimental Psychology, 56A*, 445-467.

Chen, J.-Y., Chen, T.-M., & Dell, G. S. (2002). Word-form encoding in Mandarin Chinese as assessed by the implicit priming task. *Journal of Memory and Language, 46,* 751–781.

Chen, T-M., & Chen, J-Y. (2006). Morphological encoding in the production of compound words in Mandarin Chinese. *Journal of Memory and Language, 54,* 491-514.

Chéreau, C., Gaskell, M. G., & Dumay, N. (2007). Reading spoken words: Orthographic effects in auditory priming. *Cognition, 102*, 341–360.

Chinese Linguistic Data Consortium. (2003). 现代汉语通用词表 [Chinese lexicon] (CLDC-LAC-2003-001). Beijing, China: Tsing- hua University, State Key Laboratory of Intelligent Technology and Systems, and Chinese Academy of Sciences, Institute of Automation.

Cutler, A., & Davis, C. (2012). An orthographic effect in phoneme processing, and its limitations. *Frontiers in Psychology, 3,* 18.

Damian, M. F., & Bowers, J. S. (2003). Effects of orthography on speech production in a form preparation paradigm. *Journal of Memory and Language, 49,* 119–132.

Damian, M. F., & Bowers, J. S. (2009). Assessing the role of orthography in speech perception and production: Evidence from picture-word interference tasks. *European Journal of Cognitive Psychology, 21*, 581–598.

Damian, M.F., & Dumay, N. (2009). Exploring phonological encoding through repeated segments. *Language and Cognitive Processes, 24,* 685–712.

Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review, 93*, 283-321.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review, 104,* 801-838.

Diependaele, K., Ziegler, J. C., & Grainger, J. (2010). Fast phonology and the Bi-modal interactive-activation model. *European Journal of Cognitive Psychology, 22,* 764–778.

Dijkstra, T., Roelofs, A., & Fieuws, S. (1995). Orthographic effects on phoneme monitoring. *Canadian Journal of Experimental Psychology, 49,* 264-271.

Forster, K. I., & Forster, J. C. (2003). DMDX: AWindows display pro- gram with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35,* 116–124.

Franck, J., Bowers, J. S., Frauenfelder, U. H., & Vigliocco, G. (2003). Orthographic influences on agreement: A case for modality-specific form effects on grammatical encoding. *Language and Cognitive Processes, 18*(1), 61–79.

Gaskell, M. G., Cox, H., Foley, K., Grieve, H., & O'Brien, R. (2003). Constraints on definite article alternation in speech production: To "thee" or not to "thee"? *Memory and Cognition, 31,* 715–727.

Goswami, U., Ziegler, J. C., & Richardson, U. (2005). The effects of spelling consistency on phonological awareness: a comparison of English and German. *Journal of Experimental Child Psychology, 92,* 345–365.

Grainger, J. & Ferrand, L. (1994). Phonology and orthography in visual word recognition: Effects of masked homophone primes. *Journal of Memory and Language, 33,* 218-233.

Han, J. I., & Choi, T. H. (2015). The influence of spelling on the production and storage of words with allophonic variants of /h /in Korean. *Applied Psycholinguistics*, DOI: http://dx.doi.org/10.1017/S0142716415000235

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59,* 434–446.

Kureta, Y., Fushimi, T., & Tatsumi, I. F. (2006). The functional unit in phonological encoding: Evidence for moraic representation in native Japanese speakers. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 1102–1119.

Kureta, Y., Fushimi, T., Sakuma, N., & Tatsumi, I. F. (2015). Orthographic influences on the word onset phoneme preparation effect in native Japanese speakers: Evidence from the word form preparation paradigm. *Japanese Psychological Research, 57,* 50-60.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral & Brain Sciences, 22,* 1-38.

Li, C., Wang, M., & Idsardi, W. (2015). The effect of orthographic form-cuing on the phonological preparation unit in spoken word production. *Memory and Cognition, 43,* 563-578.

Lupker, S. J. (1982). The role of phonetic and orthographic similarity in picture-word interference. *Canadian Journal of Psychology, 36*, 349–367.

Meyer, A. S. (1990). The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory and Language, 29,* 524–545.

Meyer, A. S. (1991). The time course of phonological encoding in language production: Phonological encoding inside a syllable. *Journal of Memory and Language, 30,* 69–89.

Mitterer, H., & Reinisch, E. (2015). Letters don't matter: No effect of orthography on the perception of conversational speech. *Journal of Memory and Language, 85*, 116-134.

Montant, M., Schön, D., Anton, J.-L., & Ziegler, J. C. (2011). Orthographic contamination of Broca's area. *Frontiers in Psychology, 2: 378.* doi: 10.3389/fpsyg.2011.00378

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition, 7*(4), 323–331.

Naor-Raz, G., Tarr, M. J., & Kersten, D. (2003). Is color an intrinsic property of object representation? *Perception, 32,* 667–680.

Pattamadilok, C., Perre, L., Dufau, S., & Ziegler, J. C. (2009). On-line orthographic influences on spoken language in a semantic task. *Journal of Cognitive Neuroscience, 21*, 169–179.

Perre, L., & Ziegler, J. C. (2008). On-line activation of orthography in spoken word recognition. *Brain Research, 1188,* 132–138.

Perre, L., Pattamadilok, C., Montant, M., & Ziegler, J. C. (2009). Orthographic effects in spoken language: On-line activation or phonological restructuring? *Brain Research, 1275,* 73–80.

Posnansky, C. J., & Rayner, K. (1978). Visual vs. phonemic contributions to importance of initial letter in word identification. *Bulletin of the Psychonomic Society, 11*, 188–190.

Protopapas, A. (2007). CheckVocal: A programto facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods, 39,* 859–862.

Qu, Q., Damian, M. F., & Kazanina N. (2012). Sound-sized segments are significant for Mandarin speakers. *Proceedings of the National Academy of Sciences, 109,* 14265–14270.

Rastle, K., McCormick, S. F., Bayliss, L., & Davis, C. J. (2011). Orthography influences the perception and production of speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 1588–1594.

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition, 64,* 249–284.

Roelofs, A. (2006). The influence of spelling on phonological encoding in word reading, object naming, and word generation. *Psychonomic Bulletin and Review, 13,* 33-37.

Roux, S., & Bonin, P. (2011). Cascaded processing in written naming: Evidence from the picture–picture interference paradigm. *Language and Cognitive Processes, 27,* 1–36.

Saletta, M., Goffman, L., & Brentari, D. (2015). Reading skill and exposure to orthography influence speech production. *Applied Psycholinguistics,* doi: 10.1017/S0142716415000053

Seidenberg, M. S., & Tannenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 546–554.

Snodgrass, J.G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 174-215.

Tanenhaus, M. K., Flanigan, H. P., & Seidenberg, M. S. (1980). Orthographic and phonological activation in auditory and visual word recognition. *Memory & Cognition, 8,* 513-520.

Underwood, G., & Briggs, P. (1984). The development of word recognition processes. *British Journal of Psychology, 75*, 243–255.

Wheeldon, L.R., & Monsell, S. (1992). The locus of repetition priming of spoken word production. *Quarterly Journal of Experimental Psychology, 44A*, 723-761.

Zecker, S. G., Tanenhaus, M. K., Alderman, L., & Siqueland, L. (1986). Lateralization of lexical codes in auditory word recognition. *Brain and Language, 29,* 372–389.

Zhang, Q., Chen, H. -C., Weekes, B. S., & Yang, Y. (2009). Independent effects of orthographic and phonological facilitation on spoken word production in Mandarin. *Language Speech, 52*, 113–126.

Zhang, Q., & Damian, M. F. (2009). The time course of semantic and orthographic encoding in Chinese word production: An Event-Related Potential study. *Brain Research, 1273*, 92-105.

Zhang, Q., & Damian, M. F. (2012). Effects of orthography on speech production in Chinese. *Journal of Psycholinguistic Research, 41*, 267–283.

Zhang, Q., & Weekes, B. S. (2009). Orthographic facilitation effects on spoken word production: Evidence from Chinese. *Language and Cognitive Processes, 24*, 1082–1096.

Zhao, H., La Heij, W., & Schiller, N. O. (2012). Orthographic and phonological facilitation in speech production: New evidence from picture naming in Chinese. *Acta Psychologica, 139*, 272–280.

Acknowledgements

We acknowledge Mengsi Wang and Aiping Ni for help in data collection. This work was supported by the National Natural Science Foundation of China, No. 31771212, Youth Innovation Promotion Association CAS, and the German Research Foundation (DFG) and the NSFC in project Crossmodal Learning, DFG TRR-169/NSFC No. 61621136008 to the first author.

Footnotes

[1]It is well known that subsyllabic (segmental) phonological overlap in the implicit priming task generates no priming in languages such as Mandarin (Chen, Chen & Dell, 2002) and Japanese (Kureta, Fushimi & Tatsumi, 2006), suggesting that such languages employ a different "proximate unit" than typical Western languages. Interestingly, two recent studies (Li, Wang & Idsardi, 2015; and Kureta, Fushimi, Sakuma, & Tatsumi, 2015) have demonstrated that whether or not subsyllabic overlap in the implicit priming tasks is obtained depends on the orthographic format in which the prompts are presented. However, this should probably not be interpreted as evidence for effects of spelling on spoken production, as conceptualised in the current context.

[2]A substantial number of studies have explored form priming in picture-word interference studies, and from a number of studies it is clear that with visually presented distractors, facilitation effects in PWI tasks arise not only from phonological, but also from orthographic overlap between target and distractor (e.g., Bi, Xu, & Caramazza, 2009; Damian & Bowers, 2009; Lupker, 1982; Posnansky & Rayner, 1978; Underwood & Briggs, 1984; Zhang, Chen, Weekes, & Yang, 2009; Zhang & Weekes, 2009; Zhao, La Heij, & Schiller, 2012). However, such orthographically based effects are not conventionally interpreted as implying that spoken production *per se* involves access to orthographic codes. Rather, printed distractor words activate a cohort of visually similar looking entries within the mental lexicon, among them the target, which is then produced faster than in the unprimed case. More instructive would be the demonstration of orthographic facilitation in PWI tasks with *spoken* distractors; however, this study has resulted in a null finding (Damian & Bowers, 2009, Experiment 2). Even more interesting would be a study with Chinese speakers because spelling is more easily dissociated from sound overlap. However, the prevalence of homophony (in Mandarin, each spoken syllable is estimated to map onto an average of 11 characters) makes it difficult to design experiments with spoken distractors.

[3]Tanenhaus, Flanigan and Seidenberg (1980) used a Stroop colour naming task of words ("bread") which were preceded by written or spoken orthographically related ("bead") or phonologically related ("bed") prime words, and found orthographic and phonological priming in both prime presentation formats. However, even though the primary task involved spoken

production, the orthographic effect in this study is probably best interpreted as mainly arising from word recognition and not from production.

[4]Chinese orthographic system is generally characterised as a stroke-radical-character hierarchy. A growing body of literature suggests that Chinese characters are automatically decomposed into subcharacter components, and reading or writing Chinese characters involves independent radical processing (e.g., Ding, Peng & Taft, 2004; Qu, Damian, Zhang, & Zhu, 2011; Zhou & Marslen-Wilson, 1999). These studies highlight the fact that radicals constitute important representational units in Chinese orthography. Therefore, in our study, we defined orthographic overlap in terms of a radical shared between colour and object names.

[5]In both experiments we conducted additional analyses in which the factor block/repetition was included, and we obtained a main effect of block/repetition (average latencies accelerated with repeated naming of the same objects), a main effect of relatedness, but crucially, no interaction. This suggests that the orthographic relatedness effect was largely stable across the experiment.

Table 1. *Summary of existing results on orthographic effects in spoken production, ordered by task and publication date.*

Task	Source	Language	Orthography?
Implicit priming	Chen, Chen & Dell (2002), Exp. 1	Mandarin	✗
	Damian & Bowers (2003)	English	✓
	Chen & Chen (2006), Exp. 1	Mandarin	✗
	Roelofs (2006), Exp. 2	Dutch	✗
	Alario, Perre, Castel & Ziegler (2007)	French	✗
	Bi, Wei, Janssen & Han (2009)	Mandarin	✗
	Zhang & Damian (2012)	Mandarin	✗
Association of novel word forms to objects	Rastle, McCormick, Bayliss & Davis (2011)	English	✓
	Bürki, Spinelli & Gaskell (2012)	French	✓
	Han & Choi (2015)	Korean	✓
	Saletta, Goffman & Brentari (2015)	English	?[1]
Persistent repetition priming of object naming	Wheeldon & Monsell (1992), Exp. 2 and 3	English	?[2]
Subject-verb agreement errors in sentence repetition	Franck, Bowers, Frauenfelder & Vigliocco (2003), Exp. 2	French	✗
Auditory shadowing of words with definite article	Gaskell, Fox, Foley et al. (2003)	English	✓
Naming of objects with homophonic names (case study)	Biedermann & Nickels (2008)	English	✗
Picture-word interference w. spoken distractors	Damian & Bowers (2009), Exp. 2	English	✗

[1]no orthographic effect in articulatory kinematics

[2]statistically weak difference in priming from homo- and heterographic homophones

Table 2. *Response latencies (in milliseconds) and error percentages (standard deviations in brackets) for Experiments 1 and 2.*

	Orthographically related	Orthographically unrelated	Difference
Experiment 1			
Latencies	970 (259)	986 (261)	+16
Errors	0.4 (6.0)	0.2 (3.9)	-0.2
Experiment 2			
Latencies	833 (214)	872 (248)	+39
Errors	5.2 (22.2)	5.5 (22.7)	+0.3

Appendix A

Materials used in Experiment 1

			Related			Unrelated		
Colour	Pinyin	English name	Object	Pinyin	English name	Object	Pinyin	English name
绿	lü4	green	线轴	xian4zhou2	cotton reel	杯子	bei1zi	cup
绿	lü4	green	绵羊	mian2yang2	sheep	枕头	zhen3tuo2	pillow
绿	lü4	green	纺车	fang3che1	spinning wheel	松鼠	song1shu3	squirrel
蓝	lan2	blue	花生	hua1sheng1	peanut	橡皮	xiang4pi2	rubber
蓝	lan2	blue	花瓶	hua1ping2	vase	梳子	shu1zi	comb
蓝	lan2	blue	苍蝇	cang1ying	fly	椅子	yi3zi	chair
橙	cheng2	orange	杯子	bei1zi	cup	花生	hua1sheng1	peanut
橙	cheng2	orange	梳子	shu1zi	comb	花瓶	hua1ping1	vase

橙	cheng2	orange	椅子	yi3zi	chair	纺车	fang3che1	spinning wheel
棕	zong1	brown	橡皮	xiang4pi2	rubber	线轴	xian4zhou2	cotton reel
棕	zong1	brown	枕头	zhen3tou2	pillow	绵羊	mian2yang2	sheep
棕	zong1	brown	松鼠	song1shu3	squirrel	苍蝇	cang1ying	fly

Appendix B

Materials used in Experiment 2

			Related			Unrelated		
Colour	Pinyin	English name	Object	Pinyin	English name	Object	Pinyin	English name
绿	lü4	green	绳子	sheng2zi	rope	杯子	bei1zi	cup
绿	lü4	green	绵羊	mian2yang2	sheep	梳子	shu1zi	comb
绿	lü4	green	纺车	fang3che1	spinning wheel	梯子	ti1zi	ladder
蓝	lan2	blue	花朵	hua1duo3	flower	橡皮	xiang4pi2	rubber
蓝	lan2	blue	花瓶	hua1ping2	vase	枕头	zhen3tou2	pillow
蓝	lan2	blue	苍蝇	cang1ying	fly	椅子	yi3zi	chair
橙	cheng2	orange	杯子	bei1zi	cup	花朵	hua1duo3	flower
橙	cheng2	orange	梳子	shu1zi	comb	花瓶	hua1ping1	vase

橙	cheng2	orange	梯子	ti1zi	ladder	苍蝇	cang1ying	fly
棕	zong1	brown	橡皮	xiang4pi2	rubber	绳子	sheng2zi	rope
棕	zong1	brown	枕头	zhen3tou2	pillow	绵羊	mian2yang2	sheep
棕	zong1	brown	椅子	yi3zi	chair	纺车	fang3che1	spinning wheel